



Thu Jun 13 09:38:55 2002

us-09-743-194-4.rst

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 12, 2002, 10:08:05 ; Search time 1790.11 Seconds
(without alignments)
3769.868 Million cell updates/sec

Title: US-09-743-194-4
Perfect score: 500
Sequence: 1 ggaggctgcttcacgagcg.....taagtttatcatcatcat 500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estnu: *
5: em_estov: *
6: em_estro: *
7: em_esttc: *
8: em_estti: *
9: gb_estt: *
10: gb_estt2: *
11: gb_htc: *
12: gb_gss: *
13: em_gss_hum: *
14: em_gss_inv: *
15: em_gss_pin: *
16: em_gss_vrc: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	405	81.0	470	12	AQ875364 V124F10 m
C 2	396.6	79.3	616	12	AQ502826 V74C11 mT
C 3	215	43.0	494	12	AQ875428 V125C3 mT
C 4	182.8	36.6	684	12	AQ502540 V164 mT
C 5	117	23.4	523	12	AQ875314 V122F2 mT
C 6	74.2	14.8	942	12	CNS00601
C 7	71	14.2	888	12	AZ549422 ENTBP49TF
C 8	70.2	14.0	833	12	CNS0082P
C 9	70.2	14.0	941	10	BM415213
C 10	69.6	13.9	579	12	CNS043EU
C 11	68.4	13.8	1110	12	CNS043SY
C 12	68.4	13.7	1046	12	CNS002KO
C 13	68.2	13.6	1135	12	CNS033GQ
C 14	67.8	13.6	1068	12	CNS000EV
C 15	67.6	13.5	1185	10	BF273407 GA_EB001
C 16	67.6	13.5	449	9	AL513999
C 17	67.2	13.4	694	12	AG044360 Pan trogl

18	67.2	13.4	1038	12	CNS0146B
19	67	13.4	638	12	AQ329262
C 20	67	13.4	1071	12	CNS00C6P
C 21	66.8	13.4	682	12	AG167085
C 22	66.6	13.3	431	12	CNS00YW2
C 23	66	13.2	898	12	CNS01812
C 24	66	13.2	1101	12	CNS01219
C 25	66	13.2	1101	12	CNS04XJP
C 26	65.6	13.1	859	12	AQ739567
C 27	65.6	13.1	1007	12	CNS06X9S
C 28	65.6	13.1	1885	10	BE420745
C 29	65.4	13.1	794	9	BE034370
C 30	65.4	13.1	930	10	BM415152
C 31	65.2	13.0	879	12	CNS009U5
C 32	65	13.0	1101	12	CNS00RHB
C 33	64.8	13.0	425	9	AL514791
C 34	64.8	13.0	593	12	CNS00880
C 35	64.8	13.0	797	9	BE034282
C 36	64.8	13.0	870	12	AQ866797
C 37	64.8	13.0	1101	12	CNS00EXP
C 38	64.6	12.9	815	12	AQ853920
C 39	64.4	12.9	357	12	CNS002SV
C 40	64.4	12.9	712	12	AQ036635
C 41	64.4	12.9	713	9	BE041089
C 42	64.4	12.9	858	12	AZ199419
C 43	64.4	12.9	956	12	AQ330169
C 44	64.2	12.8	388	12	CNS011JK
C 45	64.2	12.8	764	12	AZ189318

ALIGNMENTS

RESULT 1
AQ875364/c

LOCUS
DEFINITION
V124E10 mTn-3XHA/lacZ Insertion Library, strain Y2278 Saccharomyces cerevisiae genomic 5', DNA sequence.

ACCESSION
AQ875364
KEYWORDS
SOURCE

ORGANISM
baker's yeast.

REFERENCE
AUTHORS
Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A., Desbages, S.A., Cheung, K.W., Sheehan, A., Symonakis, D., Jansen, R., Umanak, L., Milgrom, R., Nelson, K., Tsakaki, R., Kanada, D., Lugo, R., Rager, J., Miller, R., Reeder, G.S. and Snyder, M.

TITLE
Genomic Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption

JOURNAL
COMMENT
Unpublished (1999)
Contact: Kumar A.
Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumar@yale.edu

Seq primer: GCCCTCTTCTTGGAGTAC
Class: transposon-tagged

Location/Qualifiers
1..470

Organism="Saccharomyces cerevisiae"
/strain="Y2278" - S288C background, cir(0) rho(0)"
/db_xref="taxon:4932"

/clone_lib="min-3XHA/lacZ insertion
/lab_host="E. coli"

/note="Vector: pHS56-Sal; A yeast genomic DNA library without 2 micron or mitochondrial DNA was prepared in pHS56-Sal; genomic DNA was size-fractionated (DNA of

FEATURES		Location/Qualifiers	
source		1. .684	
		/organism="Saccharomyces cerevisiae"	
		/db_xref="taxon:4932"	
		/clone_lib="mtn-3xHA/lacZ Insertion Library"	
		/lab_host="E. coli"	
		/note="Vector: pHS56-Sal; A yeast genomic DNA library (lacking mitochondrial DNA) was prepared in pHS56-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mtn-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance."	
BASE COUNT		165 a 149 c 111 g 259 t	
ORIGIN			
Query Match		36.6%; Score 182.8; DB 12; Length 684;	
Best Local Similarity		Pred. No. 5e-21; 32; Indels 2; Gaps 1;	
Matches		214; Conservative	
QY	252	ttctgtgcggttctctcct--ctctctgtgtttctgtgtgttcccccacgcgcatggc	309
DB	409	TTCTGTGCGGTTCCTAAATATACAGGATCATACGAGTAATCCCCATCACCGATGGAC	350
QY	310	ttatatgcgtatatatatagacgaggttttacgtcgaagatcatctcagtttcttga	369
DB	349	ATATATGCGGAAATATATATAGACGAGATTGACGTGGAAGATCATCTCAGTTTGTCTGA	290
QY	370	tagccttctcttattactttctgtttttaaactcattatacttagttttcttgatc	429
DB	289	TAGCTTTATATAAAATACTTTTCGTTAACTATAAACTTTAGTTTATTTTATGATC	230
QY	430	ggttttttctctatatacttaaaagtccaatcaagaacatacaaaaactcgtttat	489
DB	229	GGTTTTTCTCTGAATACTTGAAGTTCAAAATCAAAAGAACATACATAAACTACGTTTAT	170
QY	490	atcaatta 497	
DB	169	ATCAATTA 162	
RESULT	5		
LOCUS	AQ875314	523 bp	DNA linear GSS 08-NOV-1999
DEFINITION	V122F2 mtn-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces cerevisiae genomic 5', DNA sequence.		
ACCESSION	AQ875314		
VERSION	AQ875314.1	GI:6287460	
KEYWORDS	GSS.		
SOURCE	baker's yeast.		
ORGANISM	Saccharomyces cerevisiae		
REFERENCE	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.		
AUTHORS	1 (bases 1 to 523)		
	Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A., DesEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniatitis, D., Jansen, R., Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R., Hager, K., Miller, P., Roeder, G.S. and Snyder, M.		
TITLE	Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Kumar A Michael Snyder, Dept. of Mol. Cell. and Dev. Biology Yale University P.O. Box 208103, New Haven, CT 06520-8103, USA Tel: 203 432 9949 Fax: 203 432 6161 Email: anuj.kumare@yale.edu te of mtn-3xHA/lacZ insertion. Seq primer: GGCCTTCTTTTGGAGTAC Class: transposon-tagged.		
FEATURES			
source		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	
		1. .523	
		Location/Qualifiers	

Bernot.A., Fizames.C., Wincker.P., Brottier.P., Quetier.F.,
Saurin.W. and Weissenbach.J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
REFERENCE 3 (bases 1 to 1135)
Genoscope.
Direct Submission
Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.

FEATURES
source
1. .1135
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="208P24"
/clone_lib="G"
/note="Genoscope sequence ID : C0AG208DH12Sp1-end :
PUC-Orig"

BASE COUNT 863 a 65 c 43 g 124 t 40 others
ORIGIN

Query Match 13.7%; Score 68.4; DB 12; Length 1135;
Best Local Similarity 49.4%; Pred. No. 0.026;
Matches 171; Conservative 2; Mismatches 173; Indels 0; Gaps 0;
Qy 125 ttatcgtgttttatcagcgagattgtgctccagtgatttttttatcattcaatt 184
Db 523 TTTTTTTTTTTTTTTTTTAAATTTTWTTTTTTTTTTTTTTTTTTTTTTTT 464
Qy 185 aaggtttctaacataattttattttatcatttttagtaagctggttgctctgttc 244
Db 463 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 404
Qy 245 tgcgtcttctgcggttctctctctctctctctctctctctctctctctctctct 304
Db 403 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 344
Qy 305 tgggctatagcggtatataatagcagagtttttctcagcagatcatcagtttg 364
Db 343 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 284
Qy 365 ctgatagctttctactattactcttctggtttttaaaccattatactttctctt 424
Db 283 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 224
425 tgcggtttttttctctgtatacttaaaagttcaaatcaagaagaa 470
Db 223 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 178

RESULT 14
CNS00ENV/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC:
BACR29P07 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL069846
VERSION
AL069846.1 GI:4949798
KEYWORDS
GSS.
SOURCE
fruit fly.
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1068)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

COMMENT

- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source

1. .1068
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR29P07"
/note="end : 47"

BASE COUNT 434 a 72 c 89 g 234 t 239 others
ORIGIN

Query Match 13.6%; Score 68.2; DB 12; Length 1068;
Best Local Similarity 44.7%; Pred. No. 0.029;
Matches 160; Conservative 18; Mismatches 180; Indels 0; Gaps 0;
Qy 93 gttaaagctgtgtttatagggtgaccaggctttatctgttttatcatcgatggcgatt 152
Db 428 KTTTAAAGAKKKKKKKKKKKAAARAGKGGKTTTNTNTTTTTTTTTTTTTTTTTTT 369
Qy 153 tgtgcctccagtgatttttggatatccaaatgaagttctctacctaattttattttat 212
Db 368 TTTTNTT 309
Qy 213 catcttagttaatgctggttgctctgtctgtctgtctgtcggttctctctctctct 272
Db 308 TTTTNTT 249
Qy 273 tctgtttctctgtgtgtcccccacgcgcgagggcttatcgctatatatatagag 332
Db 248 TTTTNTT 189
Qy 333 cgagttttacgcagaagatcatctcagtttgctgtgatgccttctacttattacttt 392
Db 188 TTTTNTT 129
Qy 393 cgttttaacctcattatactttagttctctcttgatcggtttttctctctgatactt 450
Db 128 TTTTNTT 71

RESULT 15
BF273407
LOCUS
DEFINITION
GA_Eb0017023f Gossypium arboreum 7-10 dpa fiber library Gossypium
arboreum cDNA clone GA_Eb0017023f, mRNA sequence.
ACCESSION
BF273407
VERSION
BF273407.1 GI:11204477
KEYWORDS
EST.
SOURCE
Gossypium arboreum.
ORGANISM
Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE
1 (bases 1 to 1185)
Wing, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry
D., Wood, T.C., Leslie, A. and Wilkins, T.A.
An integrated analysis of the genetics, development, and evolution

	Query Match	13.6%	Score 67.8;	DB 10;	Length 1185;	
	Best Local Similarity	40.5%;	Pred. No. 0.032;			
	Matches 132;	Conservative	0;	Mismatches 194;	Indels	0; Gaps
	125	tttatcggtttttatcagatggcgattgtgcgtccagtcgtatttttggatatccaaatt	184			
QY	577	TTTTTTTTNNNTTT	636			
Db						
	185	aagggttcttcacctaaattttattttatcatcttttagttaaagtcgggttgcgtgtttc	244			
QY						
	637	NTTTTTTTTTTTTTNNNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	696			
Db						
	245	tgctgctttctgtcgggttccctctctctctgttcttctgtctgtcccccacgcgga	304			
QY						
	697	TTTTTTTTNNNTNNNNNTNNNNNTNNNTNNNTNNNTNNNNNNNNNNNNNTNT	756			
Db						
	305	tgsgccttatggcgctatatatagacgaggttttcaagtcgaagatcatcgaatttg	364			
QY						
	757	TNNNNNNNNNTNNNNNNNNNTTTTNNNTNNNNNNNNNNNNNNNNNNNNNTTT	816			
Db						
	365	cttgatagagcctttctactattatcttctgttttttaacctcatataacttaattctt	424			
QY						
	817	TT	876			
Db						
	425	tgatcggttttttctctgtatactt	450			
QY						
	877	TTTTTTNNNTTT	902			
Db						

search completed: June 12, 2002, 19:18:48
Job time: 33043 sec

